The Uterotrophic Bioassay Peer Review

William Owens
Procter & Gamble
EPA FACA – Raleigh, NC
Dec. 2, 2005

The Peer Review Process

Overall plan submitted to and approved by EDTA

Consisted of:

1. A Charge to the Peer Review Panel

2. A request for balance among member backgrounds (toxicology, endocrinology, in vitro validation)

3. Nominations from member countries

The Peer Review Process

There were 12 (10) members and 4 (2) observers.

Panel was provided with a Submission Package

There were 4 teleconferences (22 Sep 2003, 20 Oct 2003, 3 Dec 2003, 9 Feb 2004)

- Organizational
- Reply to Background & Protocol Questions
- Reply to Program Data & Procedure Questions
- Overall conclusions

Members drafted response, 2 teleconferences and exchanges of drafts and comments

The Topline:
The Panel did not reach consensus as to whether the uterotrophic bioassay was validated in the OECD program.

In fact, a very wide range of opinion offered

There appeared to be thee basic groups: validated, holding some reservations, not validated

For each telecon, each member & observer were requested to provide individual replies to the charge questions – these illustrate the range

Has Uterotrophic Bioassay been sufficiently evaluated and has its performance been satisfactorily characterized by the OECD validation program to support its proposed use for screening the potential of substances to act as oestrogen agonists and antagonists *in vivo?*

Member 5

Yes, over-all the Uterotrophic Assay has been adequately characterized. This is especially true for estrogen agonists but is weaker on the antagonists.

Member 6

The Uterotrophic Bioassay studies conducted within the OECD Program (as well as additional published data, reviewed in the background documentation) clearly justify its intended use for screening the potential of substances to act as oestrogen agonists and antagonists *in vivo*

Member 8

Yes. The test is detecting specifically and with a good sensitivity the tested estrogen agonists. Antagonists, especially weak antagonists were not tested, although their biological relevance is questionable. Overall, the in vivo assay measuring rat uterine weight changes is well characterized, reproducible in different labs and may be validated.

Member 1

Most aspects of the Uterotrophic Bioassay have been thoroughly evaluated: we know that when specified doses are tested, and analysis is performed by a central laboratory, most of the agonists will come up positive. Only a single proof-of-concept study was performed with an antagonist, and while the concept was proven, the actual performance of the test in identifying antagonists was not demonstrated. The bioassay protocol still needs to specify methods of setting doses, statistical methods, methods for identifying estrogen antagonists (relative doses of estrogen and the antagonist, and timing of administration), and criteria for identifying that a compound is or is not estrogenic. When these are addressed, and only when used as part of a multi-step process that involves receptor-binding studies and fertility assays, I believe this test would be ready to be used to help identify estrogenic compounds.

Member 2

For estrogen agonists, yes, on the basis of a limited set of agonists tested plus existing knowledge about the mechanism of uterotrophy. For antagonists, too few compounds have been tested to make a final statement. Remaining issues are however specificity (e.g. uterotrophic response to androgens) and sensitivity (e.g. guidance on dose-response testing, statistical methods and use of positive and negative controls).

Member 4

Overall, the Uterotrophic bioassay has been sufficiently evaluated by the OECD validation program to support its proposed use for screening the potential of substances to act as oestrogen agonists or antagonists *in vivo*. However, the use of only one negative substance, DBP, may limit the validation of the Uterotrophic bioassay. At least 2 to 3 negative substances should have been included to better quantify false positive.

Member 7

The study design of the Uterotrophic Bioassay did not assess the performance sufficiently. The defined data interpretation procedure, prediction model, and a final trial plan, which includes the objectives of the validation study, the study design as well as the data analysis, are lacking.

The number of test substances was not selected adequately to address the performance of the assay. For example: only one antagonist was tested at one time point, so solely the feasibility of the assay to screen for antagonists was shown, its reproducibility and predictive capacity cannot be assessed. Neither the specificity was adequately considered, since only one negative chemical was tested (that, according to literature, has been shown to act as a weak oestrogenic agonist). Before the uterotrophic assay could be used for regulatory purposes, its role in a testing strategy combined with other in vitro and in vivo tests should be defined.

Observer 1

No, the data derived from the validation study indicate the uterotrophic assay *is not suitable* for use as a screen for the identification of weak uterotrophic agonists. The data from the validation study *are insufficient* to make any conclusion regarding the utility of the test for the identification of uterotrophic antagonists.

*** several instances where the designers of this study did not follow … the OECD's guidance ….**

*** The scientific issues … derive from two fundamental flaws … incomplete and inappropriate statistical analysis … and unacceptably high between laboratory reproducibility of the test.**

****** **two fundamental flaws in the design … Firstly, an insufficient number of negative test substances … conduct additional testing of negative test substances … second is to utilise Monte Carlo simulations …**

*** second … the dosing schedule was specified to the participating laboratories. The specification of the dosing schedule controlled a significant source of between laboratory variability.**

**** statistical analysis is insufficient to characterise the predictive capacity …*** **the analysis presented to the PRP is incomplete … (observer then suggests the use of contingency statistics).**

Based on the information provided in the Submission Package:

a. Does this method adequately identify the potential for test substances to act *in vivo* as possible oestrogen agonists and antagonists?

Member 1
'Probably yes' for agonists, and 'uncertain' for antagonists.

Member 2
This judgment is limited again by the relatively low number of compounds tested. For the compounds that were tested, the assay gave favourable results, which is promising. But a definitive statement can only be made after more elaborate testing, with emphasis on antagonists and negative compounds.

Member 3

Possible oestrogen agonists – yes, but there is also a currently unmet need to be able to determine whether the significant increase in uterine weight observed is due to a chemical acting predominantly via an oestrogenic or androgenic (or other?) effect. Other relevant issues raised during the peer review also need to be addressed satisfactorily (e.g. additional details to be included in a final test protocol).
Possible oestrogen antagonists – theoretically yes, but the data available currently are too limited to reach a conclusion.

Member 4

Yes, the Uterotrophic bioassay is considered as a robust and rapid *in vivo* screening assay for possible oestrogen agonists/antagonists, based on the responsiveness in oestrogen sensitive tissue, however, the assay may have some limitations in its sensitivity and specificity compared to other endpoints used to determine a possible oestrogen activity e.g. epithelial cell height, gland number, uterine cell proliferation, lactoferrin protein induction, and measurement of the expression of oestrogen regulated genes or proteins. It is important to be aware of the limitation of the assay in the evaluation of the results.

Member 5

Yes, it is adequate but testing of additional antagonist compounds would have strengthened the antagonist issue.

Member 6

All published studies with the Uterotrophic Bioassay show that it is adequate to identify the potential of test substances to elicit the hormonal responses of interest, *i.e.* to act as oestrogen agonists and antagonists *in vivo.*

Member 8

See my first comment, for agonists with very good efficiency, for antagonists only one case proved.

Observer 1

No. See comments above. [note by Secretariat – see answer to the first question – Secretariat]

Member 7

This model could identify compounds that have a biological effect *in vivo* increasing the uterus weight but, a positive result in the utero-trophic assay cannot exclusively result because the tested chemical is an oestrogen agonist since other toxicological pathways can also lead to a stimulation of the uterus growth. On page 87-89 of the background review document the authors describe that a positive uterotrophic result can also occur with non-oestrogens, e.g. androgens, progestins, and growth factors. A definitive conclusion can only be drawn in combination with additional tests such as receptor binding tests.

Furthermore, the advantages of the uterotrophic tests have only partially been addressed in the validation study (metabolism of *non-active parental compounds* and toxicokinetics). The validation study is not confirming the ability of the uterotrophic test also to detect active metabolites from non or weak oestrogenic chemicals.

In order to decrease the level of false positives the authors suggest performing precursor assays such as ER binding. It is questionable whether an Uterotrophic test is still necessary if the chemicals have demonstrated to be positive *in vitro* assays. As discussed before the metabolic competence for this approach has not been proven. In addition, a positive in vitro result should not automatically lead to an additional *in vivo* experiment.

This approach is therefore questionable.

The Peer Review Outcome

Validated
Members 5, 6, and 8; Observer 2

Member 9

Holding reservations
Members 1, 2, 4, 6 and 10

Member 3

Clearly not validated
Member 7; Observer 1

What Were the Expressed Reservations?

1. A single negative (DBP) was inadequate.
2. A single, potent antagonist (ICI) was inadequate.
3. The statistics, diet phytoestrogen limits, and data quality criteria on immature uterine weight used should be incorporated into the guideline protocol.
4. Uterotrophic results should not be used alone (when response was limited but significant), and certainly its data should not be used for risk assessment.